

AMENDMENTS TO THE CLAIMS

Please cancel claims 102 to 128, without prejudice or disclaimer of subject matter, and amend claim 99, as shown below. This listing of claims replaces all prior versions and listings of claims in the application:

Listing of Claims:

1. to 2. (Cancelled)
3. (Previously Presented) The method of claim 99 further comprising:
recognizing a gesture associated with the object by analyzing changes in the
position information of the object, and
controlling the computer application based on the recognized gesture.
4. (Previously Presented) The method of claim 3 further comprising:
determining an application state of the computer application; and
using the application state in recognizing the gesture.
5. (Previously Presented) The method of claim 99 wherein the object is the user.
6. (Previously Presented) The method of claim 99 wherein the object is a part of the
user.
7. (Previously Presented) The method of claim 5 further comprising providing
feedback to the user relative to the computer application.
8. (Previously Presented) The method of claim 99 further comprising mapping the
position information from position coordinates associated with the object to screen coordinates
associated with the computer application.

9 to 10. (Cancelled)

11. (Previously Presented) The method of claim 99 further comprising:
analyzing the scene description to identify a change in position of the object; and
mapping the change in position of the object.

12 to 53. (Cancelled)

54. (Previously Presented) A stereo vision system for interfacing with an application program running on a computer, the stereo vision system comprising:

first and second video cameras arranged in an adjacent configuration and operable to produce at least first and second stereo video images; and

a processor operable to receive the first and second stereo video images and detect objects appearing in an intersecting field of view of the cameras, the processor executing a process to:

define an object detection region in three-dimensional coordinates relative to a position of the first and second video cameras;

divide the first and second stereo video images into features;

pair features of the first stereo video image with features of the second stereo video image;

generate a depth description map, the depth description map describing the position and disparity of paired features relative to the first and second stereo video images;

generate a scene description based upon the depth description map, the scene description defining a three-dimensional position for each feature;

cluster adjacent features;

crop clustered feature based upon predefined thresholds;

analyze the three-dimensional position of each clustered feature within the object detection region to determine position information of a control object; and

map the position information of the control object to a position indicator associated with an application program as the control object moves within the object detection region.

55. (Previously Presented) The stereo vision system of claim 54 wherein the process selects as the control object a detected object appearing closest to the video cameras and within the object detection region.

56. (Original) The stereo vision system of claim 54 wherein the control object is a human hand.

57. (Original) The stereo vision system of claim 54 wherein a horizontal position of the control object relative to the video cameras is mapped to an x-axis screen coordinate of the position indicator.

58. (Original) The stereo vision system of claim 54 wherein a vertical position of the control object relative to the video cameras is mapped to a y-axis screen coordinate of the position indicator.

59. (Original) The stereo vision system of claim 54 wherein the processor is configured to:

map a horizontal position of the control object relative to the video cameras to a x-axis screen coordinate of the position indicator;

map a vertical position of the control object relative to the video cameras to a y-axis screen coordinate of the position indicator; and

emulate a mouse function using the combined x-axis and y-axis screen coordinates provided to the application program.

60. (Original) The stereo vision system of claim 59 wherein the processor is further configured to emulate buttons of a mouse using gestures derived from the motion of the object position.

61. (Original) The stereo vision system of claim 59 wherein the processor is further configured to emulate buttons of a mouse based upon a sustained position of the control object in any position within the object detection region for a predetermined time period.

62. (Original) The stereo vision system of claim 59 wherein the processor is further configured to emulate buttons of a mouse based upon a position of the position indicator being sustained within the bounds of an interactive display region for a predetermined time period.

63. (Original) The stereo vision system of claim 54 wherein the processor is further configured to map a z-axis depth position of the control object relative to the video cameras to a virtual z-axis screen coordinate of the position indicator.

64. (Original) The stereo vision system of claim 54 wherein the processor is further configured to:

map a x-axis position of the control object relative to the video cameras to an x-axis screen coordinate of the position indicator;

map a y-axis position of the control object relative to the video cameras to a y-axis screen coordinate of the position indicator; and

map a z-axis depth position of the control object relative to the video cameras to a virtual z-axis screen coordinate of the position indicator.

65. (Original) The stereo vision system of claim 64 wherein a position of the position indicator being within the bounds of an interactive display region triggers an action within the application program.

66. (Original) The stereo vision system of claim 54 wherein movement of the control object along a z-axis depth position that covers a predetermined distance within a predetermined time period triggers a selection action within the application program.

67. (Original) The stereo vision system of claim 54 wherein a position of the control object being sustained in any position within the object detection region for a predetermined time period triggers a selection action within the application program.

68. to 98. (Cancelled).

99. (Currently Amended) A method of using computer vision to interface with a computer, the method comprising:

- capturing at least first and second images of a scene;
- dividing the first and second images into features;
- pairing features of the first image with features of the second image;
- generating a depth description map, the depth description map describing the position and disparity of paired features relative to the first and second images;
- generating a scene description based upon the depth description map, the scene description defining a three-dimensional position for each feature;
- clustering adjacent features;
- cropping clustered ~~feature~~ features based upon predefined thresholds;
- defining an object detection region;
- analyzing the three-dimensional position of each clustered feature within the object detection region to determine position information of an object; and
- using the position information to control a computer application.

100. (Previously Presented) The method of claim 99 wherein generating the scene description comprises generating the scene description from stereo images.

101. (Previously Presented) The method of claim 99 wherein:

generating a scene description comprises generating a scene description that includes an indication of a three-dimensional position of a feature included in a scene and an indication a shape of the feature; and

analyzing the scene description comprises analyzing the scene description including the indication of the three-dimensional position of the feature and the indication of the shape of the feature to determine position information of an object.

102. to 128. (Cancelled)